M&S Data Concepts and Terms

RPG Reference Document

11/30/00

Table of Contents

This document corresponds to the web version of the VV&A RPG Reference Document of the same name and date. It has been modified to make it suitable for printing.

Types of Data Used in Models and Simulations

The vast majority of models and simulations are critically dependent on data. The overall usefulness of any modeling and simulation (M&S) application is limited as much by the quality of the data[1] as by the quality of the model or simulation involved. Whether a model or simulation is used for analysis, training, or acquisition, the data involved in its preparation and execution should be subjected to the same kind of scrutiny as the model or simulation itself.

Data are symbolic representations of factual information to be used as a basis for reasoning, discussion, comprehension, communication, prediction, or calculation. However, although "factual" implies truth, "data" merely denotes information: the truth or falsity of data depends solely on the application. Data represent or "model" aspects of reality as defined in a specification. Like any model, data can never be absolutely correct for all purposes.

There are four basic types of data that support the design, development, execution and evaluation of a model or simulation:

reference data -- Descriptive information (metadata) about **all** the data used by the model, simulation, or federation, including data characteristics (e.g., resolution, fidelity, accuracy, completeness, relevancy, currency, appropriateness); specifications to which the data were developed or are provided; and factors describing data quality.

hard-wired data -- Data values implemented as part of the code (e.g., constants, set parameters).

Hard-wired data include the data values incorporated in the algorithms used to mathematically articulate the actions/reactions/interactions of the players in the mission space. Although data such as constants are included in this category, the resolution/fidelity assumptions of a simulation may require additional “facts” to be treated in this way.

instance data -- Data values comprising the baseline set of conditions (and allowable dynamic updates) under which the simulation is initiated and executed; input data (e.g., kill rates, firing ranges, flight altitudes, movement rates); and output data.

Instance data, commonly called input and output data, are data values that are stored and accessed separately from the code. They are usually found at the intersections of rows and columns in a relational database and are the facts used to initialize a simulation before it starts and to update it dynamically during execution.

---

[1] see the RPG template on Data Quality for more information

validation data -- Actual measurements from the real world or “best guess” information provided by subject matter experts (SMEs) that are used to validate that the results of the simulation are "correct enough" for the simulation to be useful. Note that validation data do not directly support the model or simulation itself, but are involved in the verification, validation, and accreditation activities.

Validation data are the real-world facts used for comparison to validate the results of a simulation. They come from empirical sources such as test ranges, live exercise results, or historical records; from outputs of other, previously validated simulations; or from the knowledge of SMEs.

A federation requires all the data needed by the individual models and simulations (i.e., federates) that comprise the federation and a fifth type:

exchange data -- All the data to be exchanged across the federation.

Exchange data are the metadata that identify what is to be exchanged in a federation. They are identified using the Object Model Template (OMT) and captured in Simulation Object Models (SOMs) and Federation Object Models (FOMs).

See the RPG reference document, *DoD Data VV&C Tiger Team White Paper* for additional information.

Metadata

To be correctly used and interpreted, data must be associated with a context. Contextual information (e.g., **Title:** name of item; **Description:** narrative definition of item; **Source:** originator of the data item) is data about data and is frequently called metadata.

metadata -- Data that describe other data or aspects of other data, such as definition, security, classification, accuracy, data type, precision, currency, source, limitations of use, and effective dates.

Metadata are used by

- M&S Users to evaluate the appropriateness of the data for their applications
- Data producers and data warehouses to evaluate the results of their processes used to generate and/or transform the data
- Data owners to assess the overall quality, utility, and value of the data for a given purpose or set of purposes

Although the amount of such contextual information that can be associated with each data value is potentially unlimited, the DoD data and M&S communities have categorized and defined types of metadata that are most appropriate for modeling and simulation needs.

Three basic types of metadata can be used to address individual data values, data elements, and entire databases:

administrative -- Information regarding distribution, points of contact, availability.

descriptive -- Information about what the data item is meant to represent, data lineage, configuration management, and traceability concerns.

quality -- Aspects of the data item that help describe or determine its accuracy, precision, completeness, logical consistency, currency, flexibility, portability, etc., and information regarding quality assessment techniques and results.

For additional information on data quality, see the RPG templates on Data Quality and the DoD Guidelines on Data Quality Management.

References

Annex C, "Data Verification, Validation, and Certification," IEEE 1278.4, *Recommended Practice for Distributed Interactive Simulation -- Verification, Validation, and Accreditation*, 1997.

DoD 8320.1-M: *Data Administration Procedures*, OASD/C3I, DTIC, Alexandria VA, 1994.

DoD 8320.1-M-1: *Data Element Standardization Procedures*, OASD/C3I, DTIC, Alexandria VA, January, 1993.

DoD 8320.1-M-3: *Data Quality Assurance Procedures*, (draft), OASD/C3I, DTIC, Alexandria, VA, February, 1994.

Rothenberg, Jeff, "A Discussion of Data Quality for Verification, Validation, and Certification (VV&C) of Data to be Used in Modeling*," Rand Project Memorandum PM-709-DMSO*, Rand, August, 1997. This is an essential guide on data quality assessment and data V&V. Includes considerations for metadata used in judging data quality and supporting data V&V.

External Links in This Document

DoD Guidelines on Data Quality Management, http://ssed1.ncr.disa.mil/srp/datadmn/dqpaper.html.

RPG References in This Document

select menu*: RPG Reference Documents*, select item: *DoD Data VV&C Tiger Team White Paper*, [also available at http://www.dmso.mil)]

select menu*: RPG Templates*, select item: "Data Quality Templates"

The appearance of hyperlinks does not constitute endorsement by the DoD, DMSO, the administrators of this website, or the information, products or services contained therein. For other than authorized activities such as military exchanges and Morale, Welfare and Recreation sites, the DoD does not exercise any editorial control over the information you may find at these locations. Such links are provided consistent with the stated purpose of this DMSO website.

§ § § § § § §